Protein aggregation and soluble aggregate formation screened by a fast microdialysis assay

Verónica Toledo-Rubio[1,2,†], Esther Vazquez [2,1,†], Gemma Platas [2,1], Joan Domingo-Espín [2,1], Ugutz Unzueta [2,1], Elke Steinkamp [2,1,], Elena García-Fruitós [2,1], Neus Ferrer-Miralles [2,1], Antonio Villaverde [2,1]*

[1] CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Spain.
[2] Institute for Biotechnology and Biomedicine and Department of Genetics and Microbiology, Autonomous University of Barcelona, Barcelona, Spain

*Corresponding author. Phone: 34 935813086; Fax: 34 935812011; E-mail: Antoni.Villaverde@uab.es

[†] VT and EV have equally contributed to this work.
Present address: Department of Molecular Medicine, Laboratory group Norbert Brattig, Bernhard-Nocht-Institute for Tropical Medicine, Bernhard-Nocht-Str. 74 20359 Hamburg, Germany

Keywords: protein aggregation; nanoparticles; dynamic light scattering; proto-aggregates; soluble aggregates

Fast screening of protein aggregation

Post-print of: Toledo-Rubio, V., et al "Protein aggregation and soluble aggregate formation screened by a fast microdialysis assay" in Journal of biomolecular screening (SAGE), vol. 15, issue 4 (2010), p. 453-457. The final version is available at DOI 10.1177/1087057110363911

Abstract

Protein aggregation is a major obstacle in recombinant protein production as it reduces the yield of soluble polypeptides. Also, the formation of aggregates occurring in the soluble fraction is more common than formerly expected, and the prevalence of these entities might significantly affect the average quality of the soluble protein species. Usually, the formation of soluble aggregates remains unperceived, since analytical methods such as dynamic light scattering are not routinely applied as quality control procedures. We have developed a methodologically simple and fast procedure, based on microdialysis and image processing, that reveals the aggregation tendency of a given protein in a specific environment. Since we also show a good correlation between macroscopic aggregation and soluble aggregate formation, the microdialysis approach also permits to estimate the occurrence of soluble aggregates.

Introduction

The conformational quality of recombinant proteins is an important matter of concern as it might dramatically influence the results of interactomic assays and other kind of proteomic studies [1]. Despite the development of useful prospective algorithms [2, 3], the aggregation tendency of polypeptides obtained by recombinant DNA procedures remains uncertain before production. In addition, protein misfolding and aggregation are observed in all the microorganisms used for recombinant protein production [4]. Bacteria, the most common protein production platform, often yields recombinant polypeptides totally or partially deposited as inclusion bodies [5]. Although recent studies have shown that inclusion bodies might contain functional protein species suitable for in situ enzymatic reactions [6-12], protein production processes are in general aimed to soluble species. Therefore, inclusion bodies are separated by differential sedimentation and further discarded. However, protein species occurring in the soluble cell fraction can form soluble aggregates [11], namely oligomeric protein clusters with globular and fibril-like morphologies [13, 14]. The formation of these structures, historically neglected, might compromise the average quality and biological activity of the soluble protein species [11, 14]. Due to the small size of soluble aggregates [13, 14], their formation is unperceived and commonly unsuspected. Also, no quality controls are routinely applied in the context of recombinant protein production to monitor the conformational quality of the soluble protein fraction.

We have here developed a methodologically simple analytical procedure that qualitatively determines the aggregation tendency of a given protein. Since intriguingly, protein deposition correlates with the size of soluble aggregates, we propose microdialysis as a routine screening method of both protein solubility and conformational quality of the soluble version under different experimental conditions.

Materials and methods

Protein production and purification

R9-GFP-H6 is a modular eGFP[15] version of 30 kDa containing cationic peptides at both amino and carboxy ends, namely nine arginines (R9) and six histidines (H6) respectively. The chimerical gene was obtained by standard cloning procedures in which two complementary synthetic oligonucleotides from TIB MOLBIOL SL, encoding the R9 cationic peptide were inserted in a pET-21b(+)-derived GFP-H6-encoding vector at the 5' end of the *gfp* gene. Precise details of this construction will be given elsewhere. R9-GFP-H6 was produced in Rosetta BL21 (DE3) *Escherichia coli* cells under the control of the T7 promoter. NLSCtHis is a recombinant β-galactosidase derived from NLSCt [16] in which a his-tag was added at the C-terminal end of the NLSCt recombinant gene. This was achieved by inserting synthetic oligonucleotides from TIB MOLBIOL SL, coding for six histidines at the 3' end of the NLSCt gene, in the pTNLSCt vector. NLSCtHis was produced in *Escherichia coli* BL21(DE3). Bovine serum albumin (BSA, ref 10735078001) was purchased from Roche.

Bacteria were cultured in 750 ml of Luria-Bertani (LB) medium (10 g/l tryptone Scharlau 07-489, 10g/l sodium chloride Scharlau SO0227, 5 g/l yeast extract Scharlau 07-079) at 37 ºC in shaker flasks until an OD=0.5 was reached, and gene expression was then induced by adding 1 mM IPTG (Applichem A4773.0005). After 3 hours, cells were harvested by centrifugation (7,650 g for 10 min at 4ºC), washed once in phosphate buffered saline (PBS) and stored at -80ºC. The pellet was resuspended in buffer A (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 20 mM Imidazole) and cells were disrupted by sonication according to standard procedures [17] in presence of a tablet of EDTA-free protease inhibitor cocktail (Complete, 11873580001 from Roche). The soluble cell fraction was separated from inclusion bodies by centrifugation at 14,841g, for 15 min at 4ºC. Upon filtration through 0.22 µm-filters, R9-GFP-H6 was purified by chromatography in Ni^{2+} columns (HiTrap Chelating HP 1 ml, 17-0408-01 GE Healthcare) in an AKTA FPLC. Positive pooled fractions in elution buffer (Tris-HCl 20 mM pH 7.5, 150 mM NaCl, 500 mM Imidazole) were collected, quantified by Bradford's procedure and stored at 4ºC.

Microdialysis

Several drops (20 µl each) of model proteins at 0.1 mg/ml in elution buffer were deposited on VSWP02500 Millipore membrane filters floating on 25 ml of each of the buffers described in Table 1, for 30 min. Dialyzed drops were collected and centrifuged, and the soluble fractions were used in dynamic light scattering (DLS) analysis.

In situ determination of fluorescence and total protein amounts

After removal of dialyzed protein membranes were further air dried. Aggregated R9-GFP-H6 deposited on them was determined by measuring direct fluorescence in a Versa Doc Imaging System 4000MP (BioRad) or after Coomasie or Ponceau staining, by using the GS-800 Calibrated Densitometer (BioRad). Data was obtained in triplicate as relative density units/mm^2 using the Quantity One program.

Dynamic light scattering

Volume size distributions of aggregates in the soluble fraction were measured using a DLS analyzer at the wavelength of 633 nm, combined with non-invasive backscatter technology (NIBS) (Zetasizer Nano ZS,Malvern Instruments Limited, Malvern, U.K.). Particle size dispersions of R9GFPH6 in different buffers were measured at 22 ºC after centrifugation (at 11,148 g for 15 min).

Statistical analysis

Linear regression analyses were performed by using the SigmaPlot software (ver 10.0), from which the obtained confidence levels (*p*) are shown.

Results and discussion

In our routine laboratory practice, we had commonly noted protein deposition on membrane filters when performing drop microdialysis, and we wondered if this fact could be related with the extent of protein aggregation in a given buffer. To evaluate this possibility, we explored the potential connection between protein deposition and aggregation by using an aggregation prone fluorescent protein (R9-GFP-H6) as a model. For that, upon production in *E. coli*, drops of freshly purified R9-GFP-H6 were microdialyzed (Figure 1A) against six different buffers commonly used for protein storage and handling (Table 1). The dialyzed material was centrifuged at 14,841 g and the protein remaining in the soluble fraction was quantified to determine the extent of aggregation, which ranged from undetectable levels to essentially 100 % of the total protein (Table 1). This indicated that buffer composition dramatically influences R9-GFP-H6 solubility in an extremely wide range, and pointed out buffer selection as a critical issue regarding solubility of aggregation-prone proteins. Interestingly, the low salt buffers 1 and 3 were the most convenient to keep R9-GFP-H6 in a soluble status.

To test the value of microdialysis as a convenient screening method for protein aggregation, the amounts of R9-GFP-H6 protein deposited on the membrane dialysis were determined indirectly by fluorescence emission, and directly by Coomassie-blue and Ponceau staining methods (Figure 1A). The visual appearance on the unstained and stained membranes was highly consistent in different aliquots of the same protein sample but dissimilar when comparing different buffers (Figure 1B). Fluorescence emission and Coomassie-blue staining values correlated, at acceptable levels of confidence, with the fraction of insoluble R9-GFP-H6 (p= 0.0518 and p= 0.0576 respectively) found in the respective buffers (Table 2). In fact, fluorescence emission and Coomassie-blue staining data evolved in parallel in different samples showing a good lineal correlation between them (p= 0.0136, not shown). This indicated a sufficient robustness in the determination of protein amounts deposited on the filters. However, protein amounts determined by Ponceau staining resulted less reliable as indicators of protein aggregation (p= 0.3521, Table 2), indicating that

in our hands, Ponceau measurements are less quantitative regarding protein amounts.

The good correlation between protein amounts deposited on the filters (determined by Coomassie blue staining) and the aggregation propensity was confirmed by using two additional (non fluorescent) proteins structurally differing from GFP, namely the extensively engineered *E. coli* β-galactosidase NLSCtHis, an structurally complex tetrameric protein (p= 0.0113), and a wild type commercial BSA, an essentially soluble protein (p= 0.0418, Table 2). This fact prompts to consider microdialysis as a reliable method for the fast screening of protein solubility, as its applicability is irrespective of the nature of the tested protein.

At this stage, we wondered if the occurrence of macroscopic aggregation could be linked to the formation of soluble aggregates, a possibility that, as far as we know, had not been so far explored. In this regard, we explored by DLS the eventual presence of soluble R9-GFP-H6 and NLSCtHis oligomers and their size. Solutions of R9-GFP-H6 recurrently showed major peaks at between 20 and 30 nm and in some cases, micro-aggregates of between 100 and 250 nm (Figure 2 A). While sizes of the smaller R9-GFP-H6 particles did not correlate with the extend of protein aggregation (p= 0.2094, Table 2), the sizes of the large particle population (> 100 nm) showed good correlation with the extent of R9-GFP-H6 in the insoluble fraction (p= 0.0086, Table 2) and therefore, also with Coomassie blue staining and fluorescence emission data (p= 0.0212 and p= 0.0422 respectively, not shown). The prevalence of the 20-30 nm particles is due to a self organized clustering of R9-GFP-H6 driven by the hanging R9 peptides (Vazquez and coauthors, accepted for publication), while the larger particles seem to represent an evolving population of protoaggregates.

On the other hand, soluble NLSCtHis peaks at around 15 nm, the size of the β-galactosidase tetramer (Figure 2B). Larger particles of between 100 and 1000 nm are observed in some buffers (Figure 2B), their size showing good correlation with the extent of protein aggregation (p= 0.0395, Table 2). The coincidence between size of soluble aggregates and the extent of insoluble protein aggregation suggest that the soluble supramolecular entities are intermediates in the protein aggregation process, leading to protein deposition as insoluble protein clusters.

In the context of the growing concerns about recombinant protein quality [1, 18], simple analytical approaches for the routine screening of protein aggregation tendency would be highly convenient. As the performance of microdialysis and further protein determination represented up around 90 min in full, and these simple procedures can be applied in parallel to a high number of samples, microdialysis is proposed here as a novel high-throughput, simple and fast prospective method appropriate for the comparative screening and monitoring of protein aggregation tendency, and indirectly, of the conformational quality of the residual population of soluble protein represented by the occurrence of soluble aggregates [14].

Acknowledgments: We are indebted to Marta Nadal and Carmen Arnan for technical assistance in DLS analysis. We appreciate the financial support received for the design and production of recombinant proteins for biomedical applications from MICINN (BIO2007-61194 and EUI2008-03610), AGAUR (2009SGR-108) and CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Spain. AV has been distinguished with an ICREA ACADEMIA award.

References

1. de Marco A: Minimal information: an urgent need to assess the functional reliability of recombinant proteins used in biological experiments. *Microb Cell Fact* 2008; 7: 20.

2. Conchillo-Solé O, de Groot NS, Avilés FX, Vendrell J, Daura X, Ventura S. AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinformatics* 2007; 8: 65.

3. Caflisch A: Computational models for the prediction of polypeptide aggregation propensity. *Curr Opin Chem Biol* 2006*;* 10: 437-444.

4. Gasser B, Saloheimo M, Rinas U, Dragosits M, Rodríguez-Carmona E, Baumann K, Giuliani M, Parrilli E, Branduardi P, Lang C, Porro D, Ferrer P, Tutino ML, Mattanovich D, Villaverde A. Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microb Cell Fact* 2008*;* 7: 11.

5. Villaverde A, Carrio MM. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol Lett* 2003; 25: 1385-1395.

6. Martinez-Alonso M, Gonzalez-Montalban N, Garcia-Fruitos E, Villaverde A. Learning about protein solubility from bacterial inclusion bodies. *Microb Cell Fact* 2009; 8: 4.

7. García-Fruitós E, Martínez-Alonso M, Gonzàlez-Montalbán N, Valli M, Mattanovich D, Villaverde A. Divergent genetic control of protein solubility and conformational quality in *Escherichia coli*. *J Mol Biol* 2007; 374: 195-205.

8. Peternel S, Grdadolnik J, Gaberc-PorekarV, Komel R. Engineering inclusion bodies for non denaturing extraction of functional proteins. *Microb Cell Fact* 2008; 7: 34.

9. Nahalka J. Physiological aggregation of maltodextrin phosphorylase from *Pyrococcus furiosus* and its application in a process of batch starch degradation to alpha-D: -glucose-1-phosphate. *J Ind Microbiol Biotechnol* 2008; 35: 219-223.

10. Nahalka J, Dib I, Nidetzky B. Encapsulation of *Trigonopsis variabilis* D-amino acid oxidase and fast comparison of the operational stabilities of free and immobilized preparations of the enzyme. *Biotechnol. Bioeng* 2008; 99: 251-260.

11. Gonzalez-Montalban N, Garcia-Fruitos E, Villaverde A. Recombinant protein solubility-does more mean better? *Nat Biotechnol* 2007; 25: 718-720.

12. Garcia-Fruitos E, Gonzalez-Montalban N, Morell M, Vera A, Ferraz RM, Aris A, Ventura S, Villaverde A. Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microb Cell Fact* 2005; 4: 27.

13. de Marco A, Schroedel A. Characterization of the aggregates formed during recombinant protein expression in bacteria. *BMC. Biochem* 2005; 6: 10.

14. Martinez-Alonso M, Gonzalez-Montalban N, Garcia-Fruitos E, Villaverde A. The functional quality of soluble recombinant polypeptides produced in *Escherichia coli* is defined by a wide conformational spectrum. *Appl Environ Microbiol* 2008; 101: 1353-1358.

15. Waldo GS, Standish BM, Berendzen J, Terwilliger TC. Rapid protein-folding assay using green fluorescent protein. *Nat Biotechnol* 1999; 17: 691-695.

16. Peluffo H, Acarin L, Aris A, Gonzalez P, Villaverde A, Castellano B, Gonzalez B. Neuroprotection from NMDA excitotoxic lesion by Cu/Zn superoxide dismutase gene delivery to the postnatal rat brain by a modular protein vector. *BMC Neurosci* 2006; 7: 35.

17. Feliu JX, Cubarsi R, Villaverde A. Optimized release of recombinant proteins by ultrasonication of *E. coli* cells. *Biotechnol Bioeng* 1998; 58: 536-540.

18. de Marco A, Sevastsyanovich YR, Cole JA. Minimal information for protein functional evaluation (MIPFE) workshop. *N. Biotechnol* 2009; 25: 170.

Figure 1

Schematic representation of the micro-dialysis procedure, shown as a sequential pattern (A). The soluble protein present in the dialyzed samples (ds) was analyzed by dynamic light scattering for comparison with fluorescence emission and protein amounts determined on the filters (Table 2). Images of micro-dialysis membranes showing the deposited R9-GFP-H6 protein dissolved in 3 different buffers (namely 1, 4 and 6; see Table 1 for composition) (B). Six replicas were processed in each filter. In V, direct visualization of untreated filters; in CBS, pictures of filter sections after Coomassie-blue staining; in DU (densitometric units), smaller sections of the same stained filters processed through the Quantity one image software.

Figure 2. Aggregate size distributions of soluble R9-GFP-H6 (A) and NLSCtHis (B) in different buffers (Table 1) measured by DLS. Plots obtained in buffers 1 (red), 4 (green) and 6 (blue) are shown as representative examples.